

User/Movie	P_1	P_2	\dots	P_M
U_1	3	4	$\dots ? \dots$?
U_2	2	?	$\dots \dots \dots$	4
U_3	4	5	$\dots \dots \dots$?
\vdots	\vdots	\vdots	\vdots	\vdots
U_N	5	1	$\dots ? \dots ? \dots$	3

User/Movie	P_1	P_2	\dots	P_M
U_1	3	4	$\dots ? \dots$?
U_2	2	?	$\dots \dots \dots$	4
U_3	4	5	$\dots \dots \dots$?
\vdots	\vdots			
U_N	5	1	$\dots ? \dots ? \dots$	3

A diagram illustrating a coordinate system. A horizontal line labeled 'M' extends from the bottom center to the right, with arrows at both ends indicating it continues infinitely. A vertical line labeled 'N' extends from the left center upwards, with arrows at both ends indicating it continues infinitely. The two lines intersect at a point representing the origin.

PREDICT

User/Movie		P_1	P_2	\dots	P_M
U_1	3	4	\dots	?	?
U_2	2	?	\dots	4	?
U_3	4	5	\dots	...	?
.	.	.	\dots
:	:	:	\dots
U_N	5	1	\dots	?	3

 M

Let's consider a subset of users and movies
and assume complete data

	Sholay	Swades	Batman	Interstellar	The Shawshank
U1	5	5	3	3	2
U2	2	4	5	5	3
U3	2	2	3	3	5
.	.	.	-	-	.
.	.	.	-	-	.
.	.	.	-	-	.

Let's consider a subset of users and movies

	Sholay	Swades	Batman	Interstellar	The Shawshank
U1	5	5	3	3	2
U2	2	4	5	5	3
U3	2	2	3	3	5
⋮	⋮	⋮	⋮	⋮	⋮

What can you say about U1?

Let's consider a subset of users and movies

	Sholay	Swades	Batman	Interstellar	The Shawshank
U1	5	5	3	3	2
U2	2	4	5	5	3
U3	2	2	3	3	5
⋮	⋮	⋮	⋮	⋮	⋮

What can you say about U1? Likes Bollywood
Dislikes Hollywood

Let's consider a subset of users and movies

	Sholay	Swades	Batman	Interstellar	The Shawshank
U1	5	5	3	3	2
U2	2	4	5	5	3
U3	2	2	3	3	5
⋮	⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮

→ What can you say about U2?

Let's consider a subset of users and movies

	Sholay	Swades	Batman	Interstellar	The Shawshank
U1	5	5	3	3	2
U2	2	4	5	5	3
U3	2	2	3	3	5
⋮	⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮

→ What can you say about U2? Likes "engineering"

Let's consider a subset of users and movies

	Sholay	Swades	Batman	Interstellar	The Shawshank
U1	5	5	3	3	2
U2	2	4	5	5	3
U3	2	2	3	3	5
.	.	.	-	-	.
.	.	.	-	-	.
.	.	.	-	-	.

→ What can you say about U3?

Let's consider a subset of users and movies

	Sholay	Swades	Batman	Interstellar	The Shawshank
U1	5	5	3	3	2
U2	2	4	5	5	3
U3	2	2	3	3	5
.	.	.	-	-	.
.	.	.	-	-	.
.	.	.	-	-	.

→ What can you say about U3? Likes "shorter" movies

	Sholay	Swades	Batman	Interstellar	Shawshank
Bollywoodness	1.2	0.95	0.01	0.01	0.02
Engineering	0.1	0.8	0.9	0.95	0.01
Length	0.1	0.08	0.2	0.25	0.90

	Sholay	Swades	Batman	Interstellar	Shawshank
Bollywoodness	1.2	0.95	0.01	0.01	0.02
Engineering	0.1	0.8	0.9	0.95	0.01
Length	0.1	0.08	0.2	0.25	0.90

* Describe each movie with some "r" features

* we have created a matrix H of size $R \times M$
 $\begin{matrix} C \\ \# \text{movies} \end{matrix}$

Bollywoodness	Engineering	Length
U ₁	4.0	0.7
V ₂
V ₃

x Describe each user with some "i" features

	Bollywoodness	Engineering	Length
U ₁	4.0	0.7	0.7
V ₂
V ₃
:			

- * Describe each user with some "n" features
- * Create a matrix W of size $N_{users} \times n_{features}$

Sholay

Swades

Batman

Interstellar

Shawshank

U1

5

5

3

3

2

U1 has rated sholay 5

Sholay

Swades

Batman

Interstellar

Shawshank

U₁

5

5

3

3

2

U₁ has rated Sholay 5

Bollywoodness

Engineering

Length

U₁

4.0

0.7

0.7

Bollywoodness

Sholay

Swades

Batman

1.2

0.95

0.01

Engineering

0.1

0.8

0.9

length

0.1

0.08

0.2

Sholay

Swades

Batman

Interstellar

Shawshank

U1

5

5

3

3

2

..

..

..

..

..

..

..

..

..

..

..

U1 has rated

Sholay

5

Bollywoodness

Engineering

Length

U1

4.0

0.7

0.7

.

:

:

:

:

:

Bollywoodness

Sholay

Swades

Batman

1.2

0.95

0.01

Engineering

0.1

0.8

0.9

length

0.1

0.08

0.2

Sholay

Swades

Batman

Interstellar

Shawshank

U1

5

5

3

3

2

U1 has rated Sholay 5

Bollywoodness

U1

4.0

Engineering

0.7

Length

0.7

Bollywoodness

Sholay
1.2

Swades
0.95

Batman
0.01

Engineering

0.1

0.8

0.9

length

0.1

0.08

0.2

Sholay

Swades

Batman

Interstellar

Shawshank

U1

5

5

3

3

2

..

..

U1 has rated Sholay 5

Bollywoodness

U1

4.0

Engineering

0.7

Length

0.7

..
:
:

Sholay
Bollywoodness

1.2

0.95

0.01

Swades
Engineering

0.1

0.8

0.9

Batman
Length

0.1

0.08

0.2

Sholay

Swades

Batman

Interstellar

Shawshank

U₁

5

5

3

3

2

U₁ has rated Sholay 5

Bollywoodness

U₁

4.0

Engineering

0.7

Length

0.7

Bollywoodness

1.2

0.95

Batman

0.01

Engineering

0.1

0.8

0.9

length

0.1

0.08

0.2

$$4 \times 1.2 + 0.7 \times 0.1 + 0.7 \times 0.1 \approx 5$$

MATRIX FACTORISATION

$$A_{N \times M} \approx W_{N \times r} H_{r \times M}$$

- * $r \ll N$ and $r \ll M$
- * Also called low rank decomposition

MATRIX FACTORISATION

$$A_{N \times M} \approx W_{N \times r} H_{r \times M}$$

* Goal: given A , learn W and H s.t.

$$A \approx WH$$

* or;

$$W^*, H^* = \underset{W, H}{\operatorname{argmin}} \left\| (A - WH) \right\|_F^2$$

Aside : NORMS

$$y = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$$

$$\|y\|_2^2 = ?$$

A side : NORMS

$$y = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$$

$$\|y\|_2^2 = ?$$

Above is square of ℓ_2 norm of y

$$= 1^2 + 2^2 + 3^2 + 4^2 = 30$$

A side : NORMS

$$A = \begin{bmatrix} 1 & 1 & 2 & 2 \\ 2 & 1 & 2 & 2 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

$$\| A \|_F^2 = ?$$

A side : NORMS

$$A = \begin{bmatrix} 1 & 1 & 2 & 2 \\ 2 & 1 & 2 & 2 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

$$\|A\|_F^2 = ?$$

↑
Frobenius norm

$$= \sum_{i=1}^3 \sum_{j=1}^4 a_{ij}^2 = 1^2 + 1^2 + 2^2 + 2^2 + 2^2 + \dots + 1^2$$

MATRIX FACTORISATION

$$W^*, H^* = \underset{W, H}{\operatorname{argmin}} \| (A - WH) \|_F^2$$

Q: How to learn W and H

MATRIX FACTORISATION

$$w^*, h^* = \underset{w, h}{\operatorname{argmin}} \| (A - wh) \|_F^2$$

METHOD I (Gradient Descent)

1) INIT w and h as $N \times r$ and $r \times N$ matrices

2) FOR i in [1, ... ITER]:

$$w = w - \alpha \frac{\partial \| A - wh \|_F^2}{\partial w}$$

$$h = h - \alpha \frac{\partial \| A - wh \|_F^2}{\partial h}$$

MATRIX FACTORISATION

$$w^*, h^* = \underset{w, h}{\operatorname{argmin}} \| (A - wh) \|_F^2$$

METHOD I (Gradient Descent)

1) INIT w and h as $N \times r$ and $r \times N$ matrices

2) FOR i in [1, ... ITER]:

$$w = w - \alpha \frac{\partial \| A - wh \|_F^2}{\partial w}$$

$$h = h - \alpha \frac{\partial \| A - wh \|_F^2}{\partial h}$$

METHOD II (Alternating least squares)

1) INIT w

2) Till convergence
- Fix w and learn h via least sq.

- Fix h and learn w via least sq.

Alternating least squares (Intro, Rest in assignment)

$$A = \begin{bmatrix} & \\ & \end{bmatrix}_{N \times M}$$
$$w = \begin{bmatrix} & \\ & \end{bmatrix}_{N \times r} \begin{bmatrix} & \\ & \end{bmatrix}^H_{n \times M}$$

Alternating least squares (Intro, Rest in assignment)

$$A = \begin{bmatrix} & \\ & \end{bmatrix}_{N \times M}$$
$$\approx$$
$$w = \begin{bmatrix} & \\ & \end{bmatrix}_{N \times r}$$
$$\begin{bmatrix} & \\ & \end{bmatrix}^n_{r \times M}$$

Remember linear regression

$$y = \begin{bmatrix} & \end{bmatrix}_{N \times 1}$$
$$\approx$$
$$x = \begin{bmatrix} & \\ & \end{bmatrix}_{N \times d}$$
$$\theta = \begin{bmatrix} & \end{bmatrix}_{d \times 1}^\top$$

$$\hat{\theta} = LS(x, y)$$

Alternating least squares (Intro, Rest in assignment)

$$A = \begin{bmatrix} \text{pink blob} \end{bmatrix}_{N \times M}$$
$$\approx$$
$$w = \begin{bmatrix} \text{light blue blob} \end{bmatrix}_{N \times r}$$
$$\theta = \begin{bmatrix} \text{green blob} \end{bmatrix}_{r \times M}$$

Remember linear regression

$$y = \begin{bmatrix} \text{pink blob} \end{bmatrix}_{N \times 1}$$
$$\approx$$
$$x = \begin{bmatrix} \text{light blue blob} \end{bmatrix}_{N \times d}$$
$$\theta = \begin{bmatrix} \text{green blob} \end{bmatrix}_{d \times 1}$$

$\hat{\theta} = LS(x, y)$

Alternating least squares (Intro, Rest in assignment)

$$A = \begin{bmatrix} \text{pink shaded block} \end{bmatrix}_{N \times M}$$
$$\approx$$
$$w = \begin{bmatrix} \text{light blue shaded block} \end{bmatrix}_{N \times r}$$
$$h = \begin{bmatrix} \text{green shaded block} \end{bmatrix}_{r \times M}$$
$$\boxed{\hat{H}[:, d] = LS(w, A[:, d])}$$

Remember linear regression

$$y = \begin{bmatrix} \text{pink shaded block} \end{bmatrix}_{N \times 1}$$
$$\approx$$
$$x = \begin{bmatrix} \text{light blue shaded block} \end{bmatrix}_{N \times d}$$
$$\theta = \begin{bmatrix} \text{green shaded block} \end{bmatrix}_{d \times 1}$$
$$\boxed{\hat{\theta} = LS(x, y)}$$